Biochemistry

Lec:3

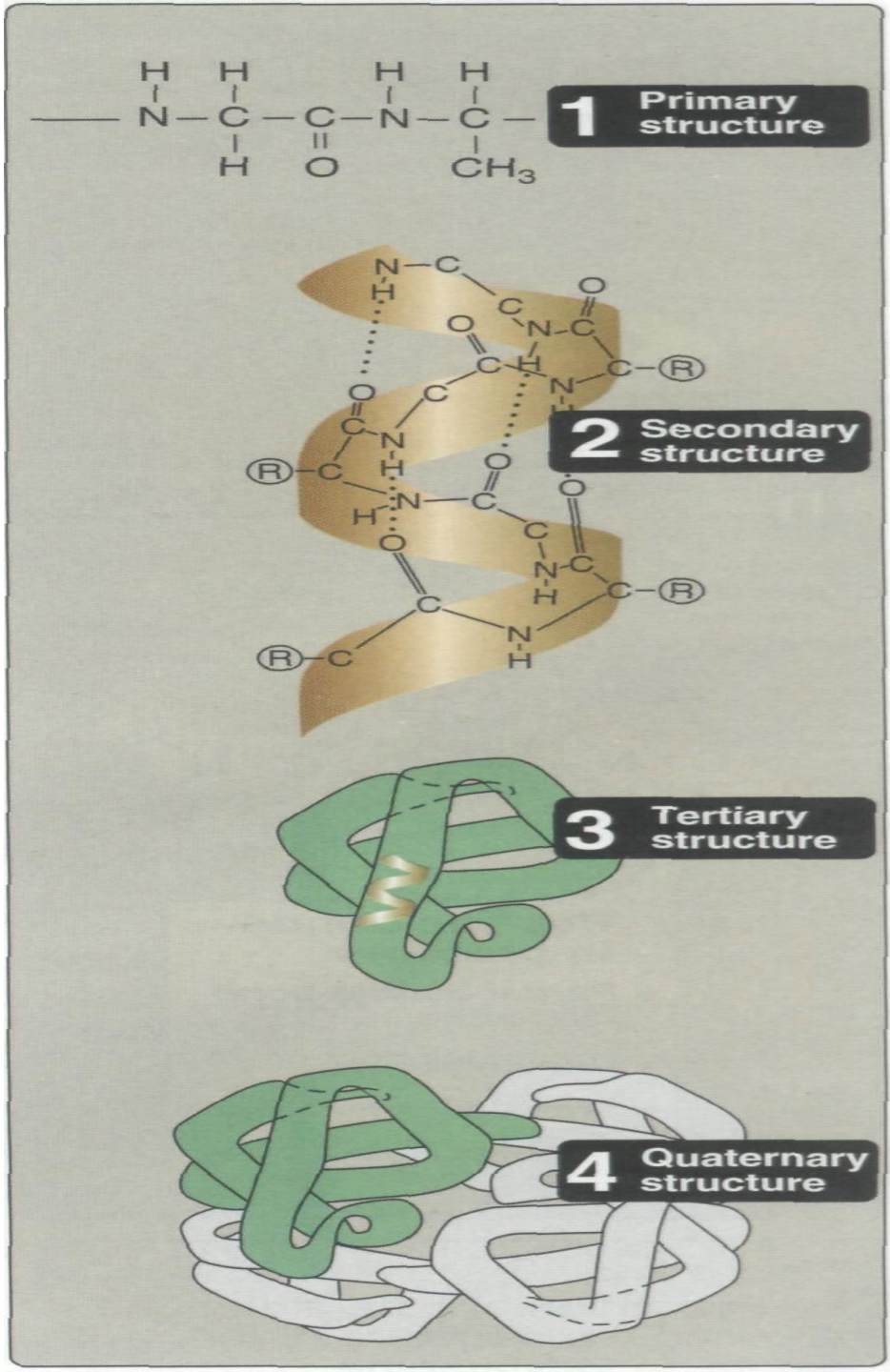
Dr.Radhwan M. Asal

Bsc. Pharmacy

MSC ,PhD Clinical Biochemistry

Structure of Proteins

The twenty amino acids commonly found in proteins are joined together by peptide bonds. The complexity of protein structure is best analyzed by considering the molecule in terms of four organizational levels, namely, primary, secondary, tertiary, and quaternary



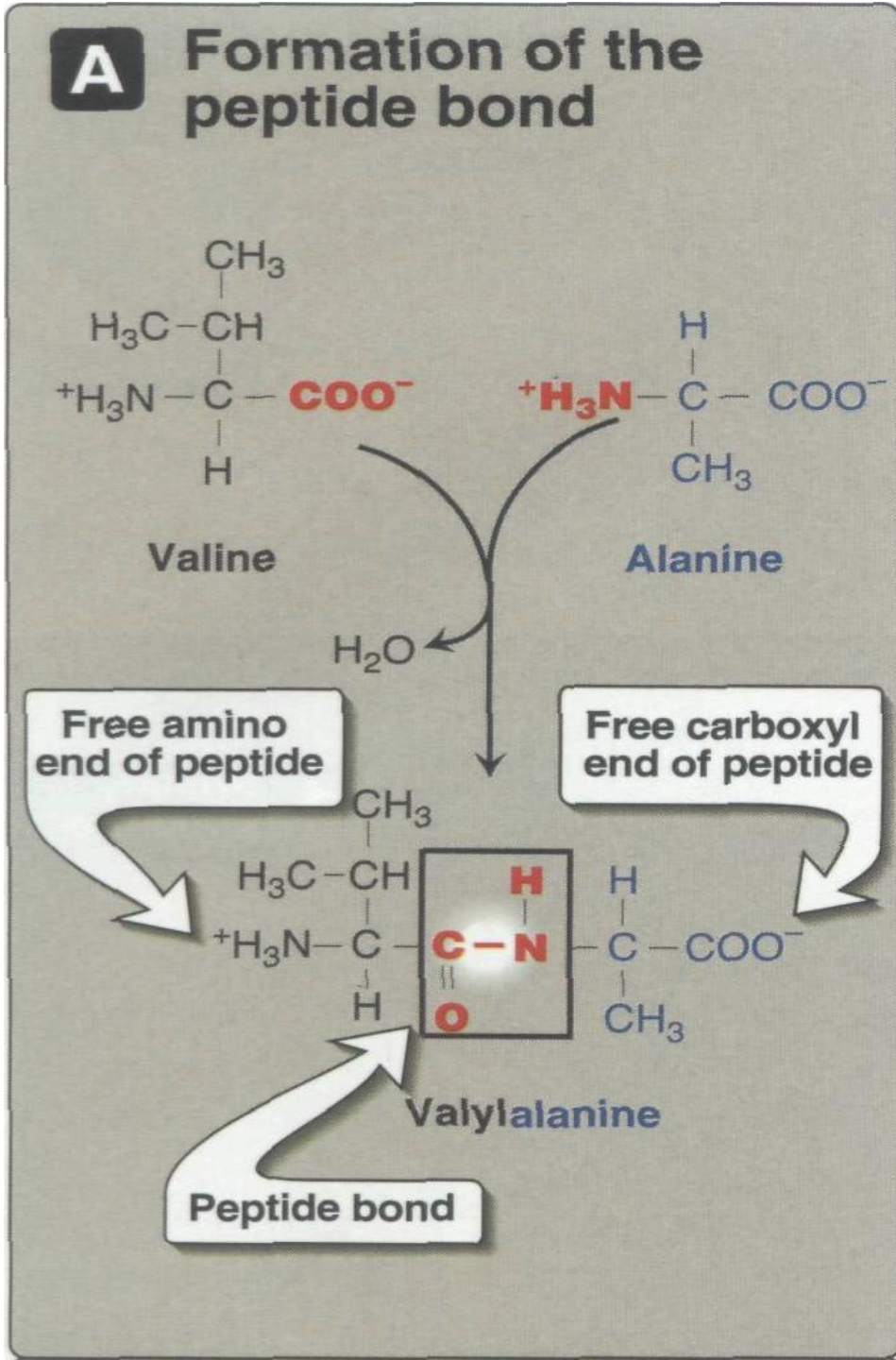
PRIMARY STRUCTURE OF PROTEINS

The **sequence of amino acids** in a protein is called the primary structure of the protein. Understanding the primary structure of proteins is important because many genetic diseases result in proteins with abnormal amino acid sequences, which cause improper folding and loss or impairment of normal function. If the primary structures of the normal and the mutated proteins are known, this information may be used to diagnose or study the disease.

Peptide bond

proteins, amino acids are joined covalently by peptide bonds, which are amide linkages between the α-carboxyl group of one amino acid, and the α-amino group of another.

For example, valine and alanine can form the dipeptide valylalanine through the formation of a peptide bond . Peptide bonds are not broken by conditions that denature proteins, such as heating or high concentrations of urea. Prolonged exposure to a strong acid or base at elevated temperatures is required to hydrolyze these bonds non enzymatically.

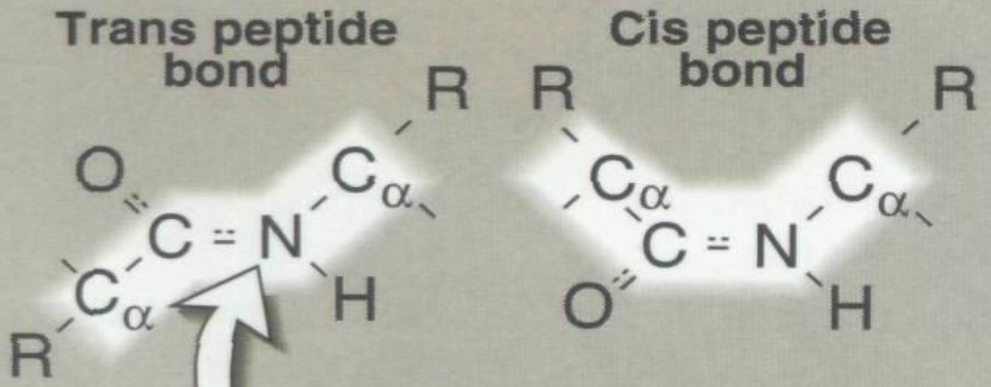


Naming the peptide: By convention, the free amino end of the peptide chain N-terminal is written to the left and the free carboxy end to the right. Therefore, all amino sequences are read from the N- to the C-terminal end of the peptide. the order of the amino acids is "valine, alanine" not "alanine, valine." Linkage of many amino acids through peptide bonds results in an unbranched chain called a **polypeptide.** Each component amino acid in a polypeptide is called a **"residue"** or **"moiety."** When a polypeptide is named, all amino acid residues have their suffixes (-ine, -an, -ic, or -ate) changed to –yl with the exception of the C-terminal amino acid. For example, a tripeptide composed of an N-terminal valine, a glycine, and a C-terminal leucine is called valylglycylleucine.

Characteristics of the peptide bond: The peptide bond has a partial double-bond character, that is, it is shorter than a single bond, and is rigid and planar . This prevents free rotation around the bond between the carbonyl carbon and the nitrogen of the peptide bond. However, the bonds between the α-carbons and the α-amino or α-carboxyl groups can be freely rotated (although they are limited by the size and character of the R-groups).

This allows the polypeptide chain to assume a variety of possible configurations. The peptide bond is generally a trans bond , in large part because of steric interference of the R-groups when in the cis position.

B Characteristics of the peptide bond



Peptide bonds in proteins

- Partial double-bond character
- Rigid and planar
- Trans configuration
- Uncharged but polar

Determination of the amino acid composition of a polypeptide

The first step in determining the primary structure of a polypeptide is to identify and quantitate its constituent amino acids. A purified sample of the polypeptide to be analyzed is first hydrolyzed by strong acid at 110 C^o for 24 hours. This treatment cleaves the peptide bonds, and releases the individual amino acids, which can be separated by **cation-exchange chromatography.** In this technique, a mixture of amino acids is applied to a column that contains a resin to which a negatively charged group is tightly attached. [Note: If the attached group is positively charged, the column becomes an **anion-exchange column.]** The amino acids bind to the column with different affinities, depending on their charges, hydrophobicity, and other characteristics.

Each amino acid is sequentially released from the chromatography column by eluting with solutions of increasing ionic strength and pH . The separated amino acids contained in the eluate from the column are quantitated by heating them with ninhydrine a reagent that forms a purple compound with most amino acids, ammonia, and amines. The amount of each amino acid is determined spectrophotometrically by measuring the amount of light absorbed by the ninhydrin derivative.

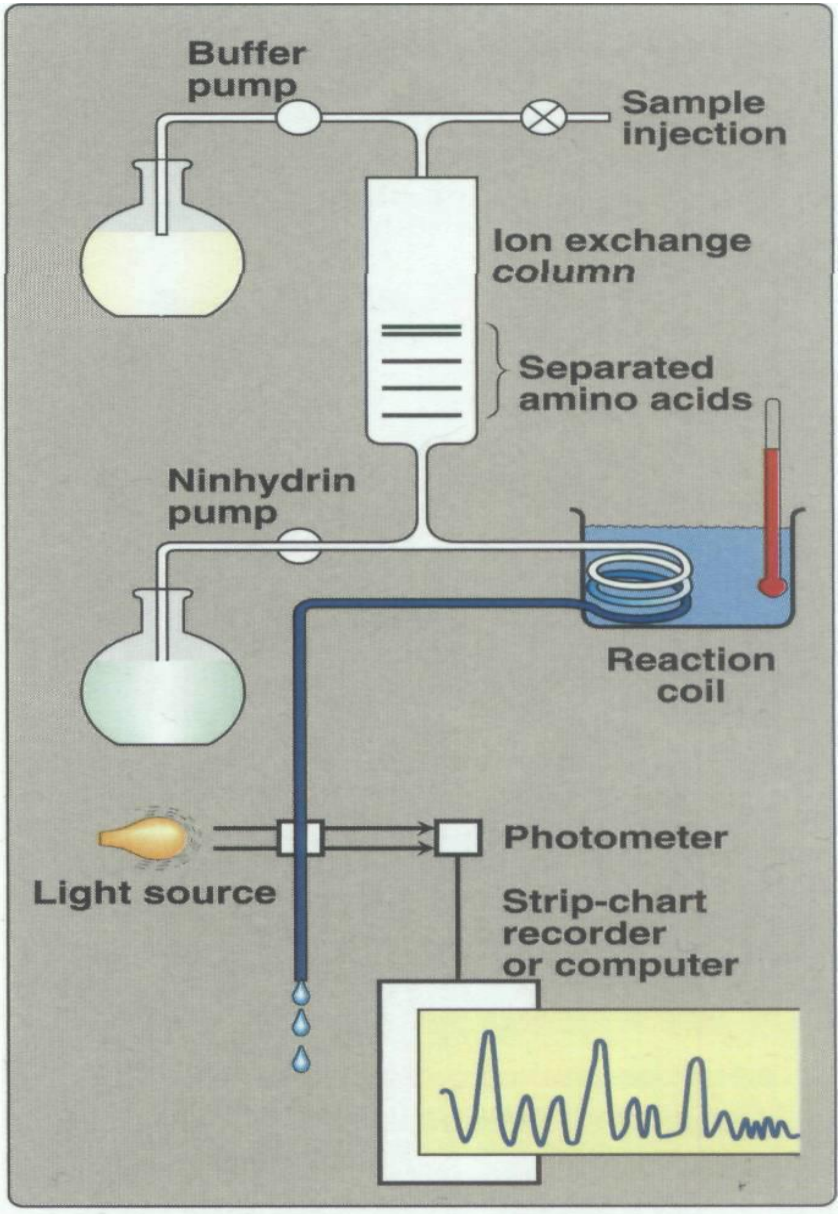


Figure 2.3
Determination of the amino acid composition of a polypeptide using an amino acid analyzer.

Sequencing of the peptide from its N-terminal end

Sequencing is a stepwise process of identifying the specific amino acids at each position in the peptide chain, beginning at the N terminal end. Phenylisothiocyanate, known as Edmans reagent, is used to label the amino terminal residue under mildly alkaline conditions .

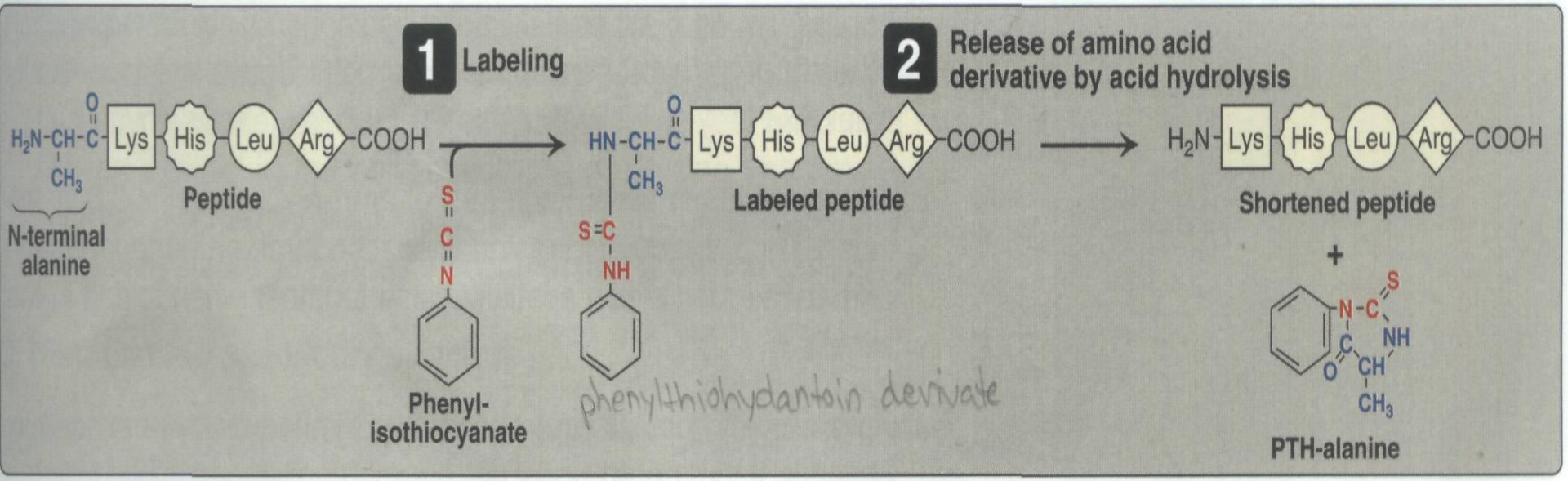


Figure 2.4
Determination of the amino-terminal residue of a polypeptide by Edman degradation.

The resulting phenylthiohydantoin (PTH) derivative introduces an instability in the N-terminal peptide bond that can be selectively hydrolyzed without cleaving the other peptide bonds. The identity of the amino acid derivative can then be determined. Edman's reagent can be applied repeatedly to the shortened peptide obtained in each previous cycle. This process has been automated and, currently, the repetition of the method can be employed by a machine (sequenator) to determine the sequence of more than 100 amino acid residues, starting at the amino terminal end of a polypeptide.

--Cleavage of the polypeptide into smaller fragments

Many polypeptides have a primary structure composed of more than 100 amino acids. Such molecules cannot be sequenced directly from end to end by a sequenator. However, these large molecules can be cleaved at specific sites, and the resulting fragments sequenced. By using more than one cleaving agent (enzymes and/or chemicals) on separate samples of the purified polypeptide, overlapping fragments can be generated that permit the proper ordering of the sequenced fragments, thus providing a complete amino acid sequence of the large polypeptide.

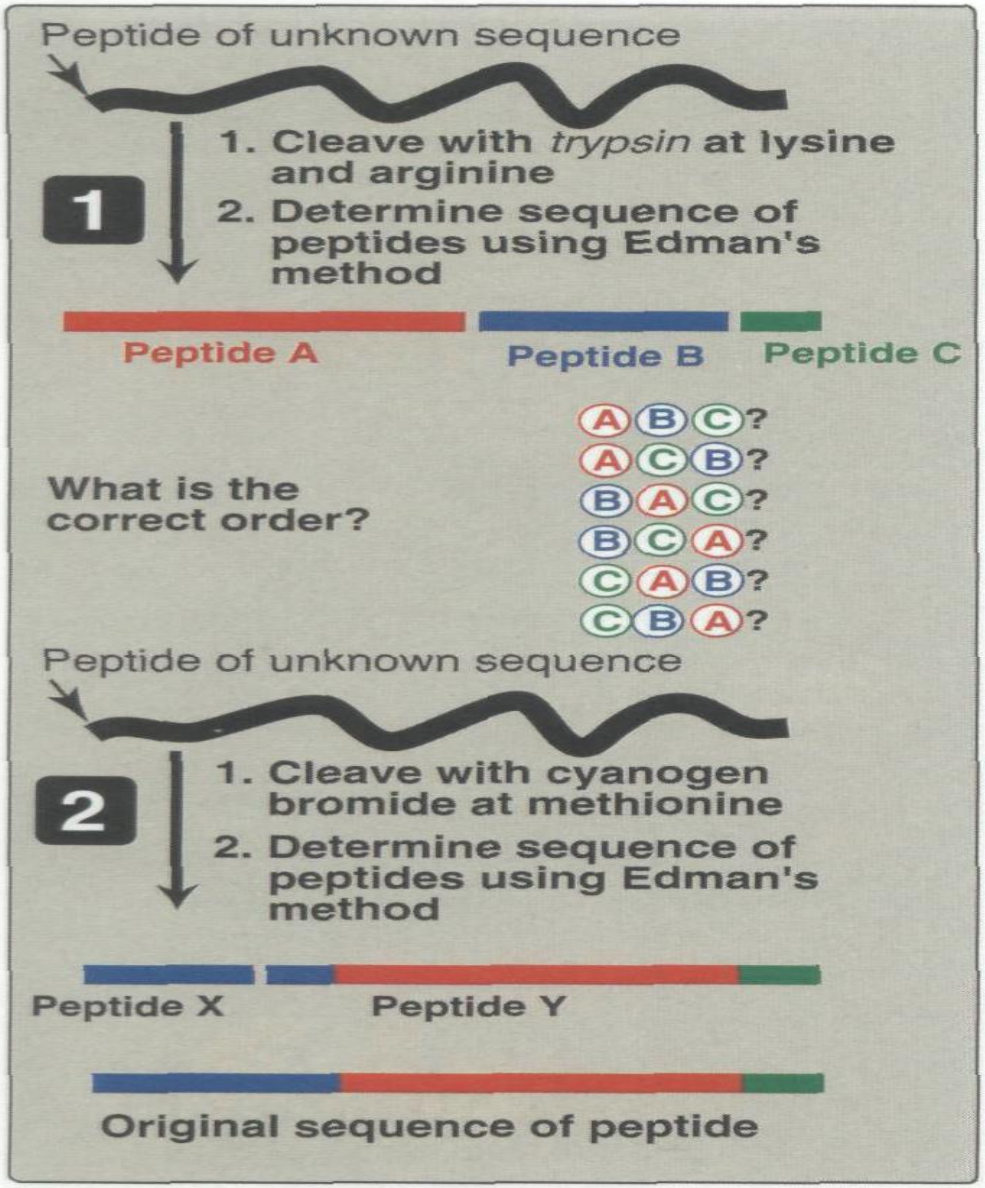


Figure 2.5
Overlapping of peptides produced by the action of *trypsin* and cyanogen bromide.

Determination of a protein's primary structure by DNA sequencing

The sequence of nucleotides in a coding region of the DNA specifies the amino acid sequence of a polypeptide. Therefore, if the nucleotide sequence can be determined, it is possible, from knowledge of the genetic code , to translate the sequence of nucleotides into the corresponding amino acid sequence of that polypeptide. This process, although routinely used to obtain the amino acid sequences of proteins, has the limitations of not being able to predict the positions of disulfide bonds in the folded chain, and not identifying any amino acids that are modified after their incorporation into the polypeptide (post-translation modification).

SECONDARY STRUCTURE OF PROTEINS

The polypeptide backbone does not assume a random three-dimensional structure, but instead generally forms regular arrangements of amino acids that are located near to each other in the linear sequence. These arrangements are termed the **secondary structure** of the polypeptide. The α-helix, β-sheet, and β-bend are examples of secondary structures frequently encountered in proteins.

α-helix

There are several different polypeptide helices found in nature, but the α-helix is the most common.

It is a spiral structure, consisting of a tightly packed, coiled polypeptide backbone core, with the side chains of the component amino acids extending outward from the central axis to avoid interfering sterically with each other .

A very diverse group of proteins contains α-helices. For example, the keratins are a family of closely related, fibrous proteins whose structure is nearly entirely α-helical. They are a major component of tissues such as hair and skin, and their rigidity is determined by the number of disulfide bonds between the constituent polypeptide chains. In contrast to keratin, myoglobin, whose structure is approximately eighty percent α-helical, is a globular, flexible molecule .

Hydrogen bonds: An α-helix is stabilized by extensive hydrogen bonding between the peptide-bond carbonyl oxygens and amide hydrogens that are part of the polypeptide backbone. The hydrogen bonds extend up the spiral from the carbonyl oxygen of one peptide bond to the -NH - group of a peptide linkage four residues ahead in the polypeptide.

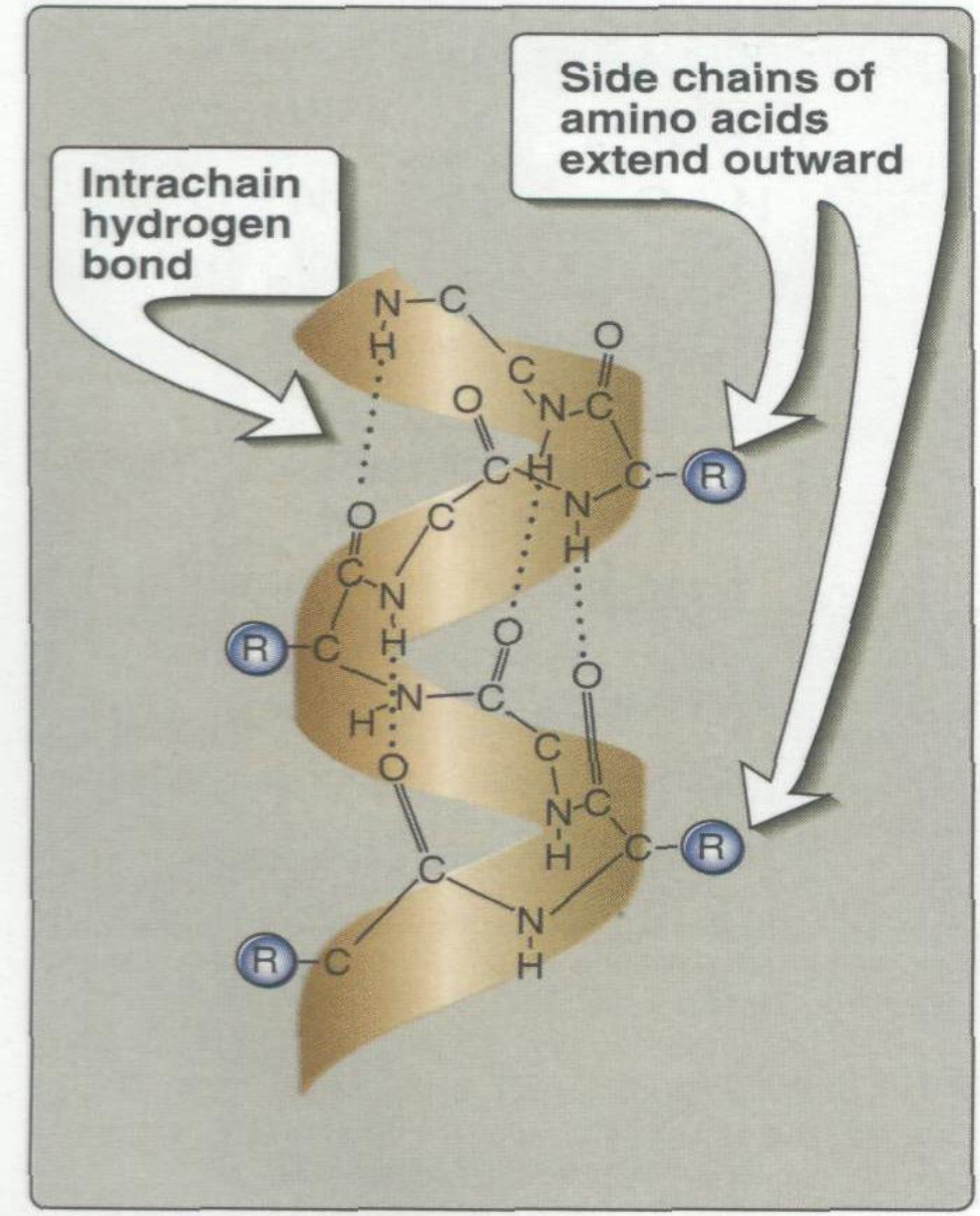


Figure 2.6
α-Helix showing peptide backbone.

This ensures that all but the first and last peptide bond components are linked to each other through hydrogen bonds. Hydrogen bonds are individually weak, but they collectively serve to stabilize the helix.

Amino acids per turn: Each turn of an α-helix contains 3.6 amino acids. Thus, amino acid residues spaced three or four apart in the primary sequence are spatially close together when folded in the α-helix.

Amino acids that disrupt an α-helix: Proline disrupts an α-helix because its imino group is not geometrically compatible with the right-handed spiral of the α-helix. Instead, it inserts a kink in the chain, which interferes with the smooth, helical structure.

Large numbers of charged amino acids (for example, glutamate aspartate, histidine, lysine, or arginine) also disrupt the helix by forming ionic bonds, or by electrostatically repelling each other. Finally , amino acids with bulky side chains, such as tryptophan, or amino acids, such as valine or isoleucine, that branch at the β-carbon(the first carbon in the R-group, next to the α-carbon) can interfere with formation of the α-helix if they are present in large numbers.

β-sheet

The β-sheet is another form of secondary structure in which all of the peptide bond components are involved in hydrogen bonding. The surfaces of β-sheets appear "pleated," and these structures are, therefore, often called **"β-pleated sheets."** When illustrations are made of protein structure, β-strands are often visualized as broad arrows .

Comparison of a β-sheet and an α-helix: Unlike the α-helix,β-sheets are composed of two or more peptide chains (β-strands),or segments of polypeptide chains, which are almost fully extended. Note also that in β-sheets the hydrogen bonds are perpendicular to the polypeptide backbone.

Parallel and antiparallel sheets: A β-sheet can be formed from two or more separate polypeptide chains or segments of polypeptide chains that are arranged either antiparallel to each other (with the C-terminal and N-terminal ends of the β-strands alternating , or parallel (with all the N terminal of the β-strands together .

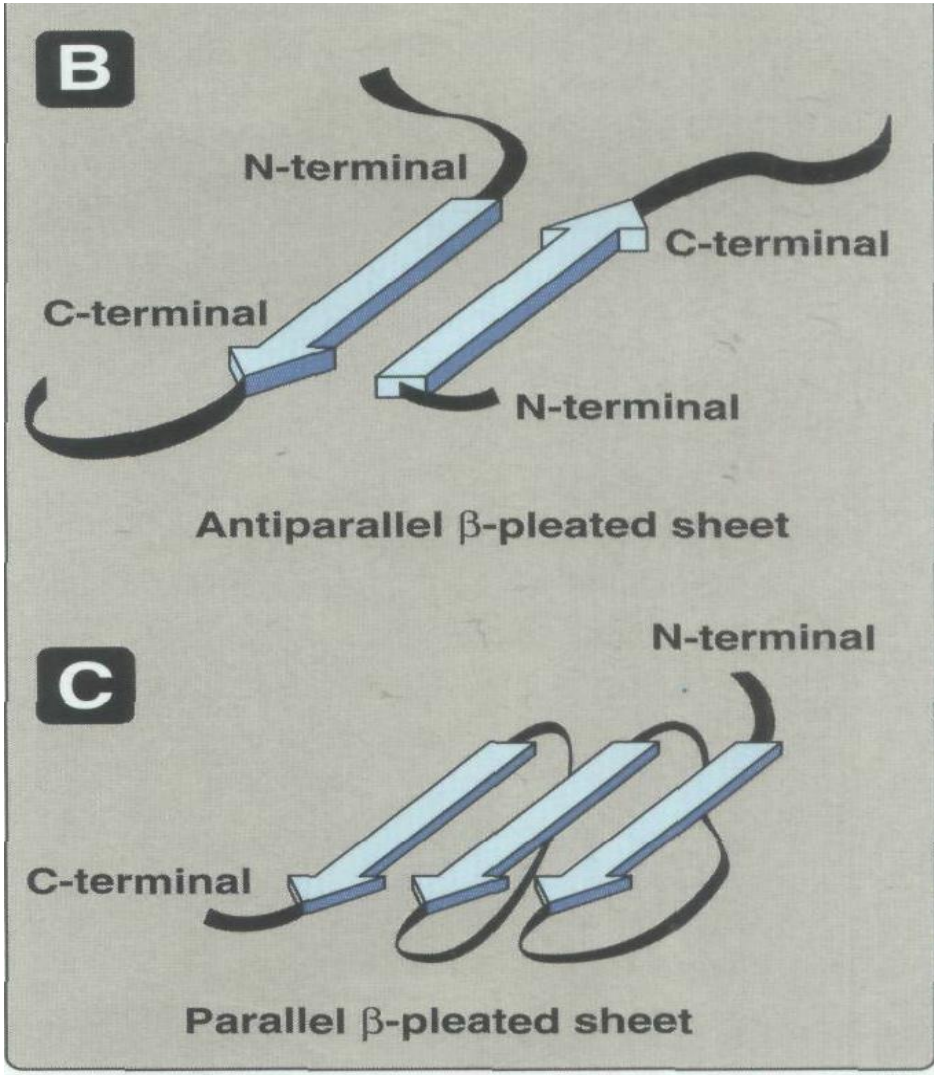


Figure 2.7
A. Structure of a β-sheet. B. An anti-parallel β-sheet with the β-strands represented as broad arrows. C. A parallel β-sheet formed from a single polypeptide chain folding back on itself.

When the hydrogen bonds are formed between the polypeptide backbones of separate polypeptide chains, they are termed **inter-chain bonds. A** β-sheet can also be formed by a single polypeptide chain folding back on itself . In this case, the hydrogen bonds are **intrachain bonds.**

globular proteins, β-sheets always have a right-handed curl, or twist, when viewed along the polypeptide backbone. [Note: Twisted β-sheets often form the core of globular proteins.]

B-bends (reverse turns)

B-bends reverse the direction of a polypeptide chain, helping it forma compact, globular shape. They are usually found on the surface of protein molecules, and often include charged residues.

[Note: B-Bends were given this name because they often connect successive strands of antiparallel β-sheets.] B-Bends are generally composed of four amino acids, one of which may be proline the imino acid that causes a "kink" in the polypeptide chain. Glycine, the amino acid with the smallest R-group, is also frequently found in β-bends. B-Bends are stabilized by the formation of hydrogen and ionic bonds.

Nonrepetitive secondary structure

Approximately one half of an average globular protein is organized into repetitive structures, such as the α-helix and/or β-sheet. The remainder of the polypeptide chain is described as having a loop or coil conformation. These nonrepetitive secondary structures are not random," but rather simply have a less regular structure than those described above.

Supersecondary structures (motifs)

Globular proteins are constructed by combining secondary structural elements (α-helices, β-sheets, nonrepetitive sequences). These form primarily the core region that is, the interior of the molecule. They are connected by loop regions (for example, β-bends) at the surface of the protein. Supersecondary structures are usually produced by packing side chains from adjacent secondary structural elements close to each other. Thus, for example, α-helices and β-sheets that are adjacent in the amino acid sequence are also usually adjacent in the final, folded protein.